Mining and predictive characterization for resistance to leaf rust (Puccinia hordei Otth) using two subsets of barley genetic resources

Mariam Amouzoune
ICARDA: International Center for Agricultural Research in the Dry Areas
Ahmed Amri
ICARDA: International Center for Agricultural Research in the Dry Areas
Rachid Benkirane
Ibn Tofail University Kenitra: Universite Ibn Tofail Kenitra
Zakaria Kehel
ICARDA: International Center for Agricultural Research in the Dry Areas
Muamer Al-Jaboobi
ICARDA: International Center for Agricultural Research in the Dry Areas
Adil Moulakat
ICARDA: International Center for Agricultural Research in the Dry Areas
Jilal Abderrazek
Institut National de Recherche pour l'Agriculture l'Alimentation et l'Environnement
Sajid Rehman (srehman@perennia.ca)
Perennia Food and Agriculture Inc.

Research Article

Keywords: Barley, Puccinia hordei, Resistance, Genetic resources, Efficient mining, Machine learning.

Posted Date: April 16th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-394214/v1

Version of Record: A version of this preprint was published at Genetic Resources and Crop Evolution on October 2nd, 2021. See the published version at https://doi.org/10.1007/s10722-021-01268-4.

Mining and predictive characterization for resistance to leaf rust (*Puccinia hordei* Otth) using two subsets of barley genetic resources

Mariam Amouzoune[1, 3], Ahmed Amri[3], Rachid Benkirane[1], Zakaria Kehel[3], Muamer Al-Jaboobi[3], Adil Moulakat[3], Jilal Abderrazek[4], and Sajid Rehman[2, 3*]

[1] Faculty of Sciences, University Ibn Tofail, Kenitra, Morocco
[2] Perennia Food and Agriculture Inc., Nova Scotia
[3] International Center for Agricultural Research in the Dry Areas (ICARDA), Rabat, Morocco
[4] Institut National de la Recherche Agronomique (INRA), Rabat, Morocco

[*]Corresponding author: Sajid Rehman
Email : srehman@perennia.ca

Mariam Amouzoune	Mariam.amouzoune@gmail.com
Ahmed Amri	a.amri@cgiar.org
Rachid Benkirane	rachid.benkirane@yahoo.fr
Zakaria Kehel	Z.Kehel@cgiar.org
Muamer Al-Jaboobi	moumr@hotmail.com
Adil Moulakat	A.Moulakat@cgiar.org
Jilal Abderrazek	jilal abderrazek_2001@yahoo.fr
Sajid Rehman	srehman@perennia.ca

Abstract

Sustainable barley production will require access to diverse *ex-situ* conserved collections to develop varieties with high yields and capable of overcoming the challenges imposed by major abiotic and biotic stresses. This study aimed at searching efficient approaches for the identification of new sources of resistance to barley leaf rust (LR). Two subsets, Generation Challenge Program Reference set (GCP) with 190 accessions and leaf rust subset constructed using the filtering approach of the Focused Identification of Germplasm Strategy (FIGS) with 100 accessions, were evaluated for the seedling as well as the adult plant stage resistance (APR) using two LR isolates (ISO-SAT and ISO-MRC) and in four environments in Morocco, respectively. Both subsets yielded a high percent of accessions with a moderately resistant (MR) reaction to the two LR isolates at the seedling stage. For APR, more than 50% of the accessions showed resistant reactions in SAT2018 and GCH2018, while this rate was less than 20% in SAT2017 and SAT2019. Statistical analysis using chi-square test of independence revealed the dependency of LR reaction on subsets at the seedling (ISO-MRC), as well as at the APR (SAT2017 and SAT2018) stage. Furthermore, the test of goodness of fit showed that FIGS_LR yielded higher percentages of resistant accessions than GCP subset in case of ISO-MRC at the seedling stage, and in case of SAT2017 and SAT2018 at APR stage. Although some of the tested machine learning models had moderate to high accuracies, none of them was able to find a strong and significant relationship between the reaction to LR and the environmental conditions showing the needs for more fine tuning of approaches for efficient mining of genetic resources using machine learning.

Keywords Barley, *Puccinia hordei*, Resistance, Genetic resources, Efficient mining, Machine learning.

Introduction

Cultivated barley (*Hordeum vulgare* subsp. *vulgare* L.) is the fourth most important cereal crop in the world after wheat, maize, and rice, in terms of production of 143.13 million metric tons and acreage of around 47.37 million hectares (FAOSTAT 2017; USDA 2019). In Morocco, barley is grown on 2 million hectares in the arid and semi-arid regions with 1.23 t/ha of average grain yield, which is relatively low compared to North America (3.67 t/h) (FAOSTAT 2017). The lower national average grain yield of barley is due to limited or no use of inputs, and the prevalence of abiotic and biotic constraints. Foliar diseases such as powdery mildew, net blotch, spot blotch, and leaf rust are important biotic constraints that limit the grain and straw yields and their quality. Barley leaf rust caused by *Puccinia hordei* Otth (*Ph*) is one of the most destructive and globally spread barley diseases (Clifford 1985; Park *et al.* 2015). It is widely distributed throughout barley growing area, and can cause serious yield losses in the regions of North Africa, Europe, New Zealand, Australia, the Eastern and Midwestern parts of United States, and some parts of Asia, where susceptible and late maturity varieties of barley are sown (Arnst *et al.* 1979; Clifford 1985; Chicaiza *et al.* 1996; Brunner *et al.* 2000, Niks *et al.* 2000). Losses of barley production due to LR can reach up to 30% on susceptible cultivars (Cotterill *et al.* 1992, Griffey *et al.* 1994).

Applying fungicides is an efficient strategy to control major foliar diseases, but it is not economical for barley grown under marginal lands and low-input conditions of Morocco. Therefore, the use of resistant varieties is the most effective, economical, and environmentally safe way for controlling barley LR. This can be achieved by transferring identified resistant genes from diverse genetic resources into elite barley germplasm (Hajjar and Hodgkin, 2007; Rehman *et al.* 2020). To date, 23 *Rph* genes conferring hypersensitive resistance to barley leaf rust at the seedling stage (*Rph1* to *Rph19*, *Rph21,* and *Rph22, Rph25, Rph26*), and 3 APR genes (*Rph20, Rph23, Rph24*) have been identified from *Hordeum vulgare* subsp. *vulgare*, or transferred from *H. vulgare* subsp. *spontaneum*, and *H. bulbosum* (Park 2015; Kavanagh *et al.* 2017; Yu *et al.* 2018). However, LR resistance faces a big challenge from a rapidly evolving pathogen due to recombination and mutations which leads to the development of new pathotypes that overcome deployed single major *Rph* genes in a short time span (McIntosh 1988; Figueroa *et al.* 2016). Most of the barley varieties released in Morocco are susceptible to LR, and limited sources of resistance are available against the LR populations prevailing in the northern parts of Morocco. Therefore, it is required to evaluate and identify continuously new sources of resistance from existing germplasm, and from gene bank collections (Qualset 1975; Sing *et al.* 2015).

The genetic resources remain the most important source of parental germplasm for barley breeding programs to develop new varieties with high yield, better end-use quality, tolerant to abiotic stresses, and resistant to major diseases and pests. But the search for a given trait is limited owing to the large number of accessions being held in the genebanks. Further, the evaluation of these large collections for some traits can be very expensive. To facilitate the screening and the mining of genetic resources, it requires the development of intelligent sub-setting approaches to fit the available funding and facilities (ICARDA 2015). These approaches aim to select subsets from the original collection to harness maximum diversity within limited number of accessions (Gollin *et al.* 2000). Frankel and Brown (1984) recommended the use of core collection which selects 5-10% of the original collection, representing maximum geographical or morphological diversities. However, because of the large number of accessions in the entire collection in genebanks, even a core collection can still be unmanageable for the evaluation of some traits, and other sub-setting approaches were suggested. Mini-core collections were

suggested by Upadhyaya and Ortiz (2001) to concentrate broad genetic diversity in smaller subsets. It allows selecting about 1% of the total accessions from the entire collection and the core collection to represent maximum diversity. The Generation Challenge Program (GCP) (https://www.generationcp.org) recommended the development of a reference set representing 10% of the core collection to represent maximum diversity using molecular markers.

From ICARDA barley in-trust collection totaling more than 32,000 accessions, the barley core collection (composite set) of 3,000 accessions of both cultivated and wild progenitor species (*H. spontaneum*) was selected based on climatological data of the collection sites; from which the Generation Challenge Program (GCP) developed a reference subset of 300 accessions based on the diversity of EST-derived, and genomic SSR markers (https://www.croptrust.org/wp/wp-content/uploads/2014/12/Barley_Strategy_FINAL_27Oct08.pdf). Though, many researchers have reported on the limitations of core collections in capturing rare and adaptive alleles (Dwivedi *et al.* 2008; Xu 2010).

The Focused Identification of Germplasm Strategy (FIGS) was developed by ICARDA in collaboration with the Australian and the Russian partners as an alternative approach for efficient mining of genetic resources that maximize the likelihood of capturing specific adaptive traits in subsets of manageable size extracted from the original collection (Mackay 1990; Street *et al.* 2008). FIGS is based on the co-evolution between the accessions and the environmental conditions in which they evolved (Mackay 1995; Gollin *et al.* 2000; Mackay and Street 2004; Bari *et al.* 2012). This approach exploits the development of the relationship between the specific sought-trait and ecogeographical data by filtering germplasm collections through exerting selection pressures of the emergence of a sought trait. It uses also the modeling pathways. When the relationship is confirmed, a manageable subset can be selected to include accessions with high probability of having the sought trait. FIGS subsets have allowed to identify for the first-time sources of resistance to Sunn pest in wheat (El Bouhssini *et al.* 2009), resistance to net blotch in barley (Endersen *et al.* 2011), and drought tolerance in faba bean (Khazaei *et al.* 2014). The present study aimed at: i) Identification of sources of resistance to LR in FIGS_LR and GCP subsets; ii) Assessing the dependence of resistance on sub-setting approach; and iii) search for the best model that describes the relationship between resistance to LR, and the environmental conditions using machine learning.

Materials and Methods

Plant material

Two barley subsets extracted from ICARDA in-trust collection available from the regeneration efforts conducted in Morocco to reconstruct the active and base collections were used in this study. A total of 188 accessions from the reference set constructed within the Generation Challenge Program (GCP) and extracted from the composite set of barley collection held at ICARDA based on diversity of EST-derived and genomic SSR markers (Supplementary Table S1). Another subset composed of 86 accessions was selected using filtering approach of the Focused Identification of Germplasm Strategy (FIGS_LR) based on the following parameters:

- ✓ Count number of days where the average daily temperature is between 8 – 15 °C, 10 days before the onset of growing period and up to 15% into the vegetative phase.
- ✓ Remove sites with zero count from step 1.

- ✓ Sum daily rain for 10 days before the onset of growing period up to 10% into the vegetative phase.
- ✓ Normalize both variables (steps 1 and 2) to range 0 -1 for each site.
- ✓ Add variables to create index 1.
- ✓ Rank based on index 1 and remove bottom 25 percent of sites.

For the remaining sites, the following was done:

- ✓ From 10% into the vegetative phase until onset of grain filling divide into 3 separate sub-phases of equal length.
- ✓ For each sub-phase count the number of days where the average daily temperature is between 18 – 20 ° C.
- ✓ For each sub-phase determine the amount of precipitation.
- ✓ Remove sites if any of the variables = 0 (3 count variables and 3 precipitation variables).
- ✓ Normalize each variable for a range between 0 – 1.
- ✓ Add each variable and then add index 1 to create index 2.
- ✓ Rank sites using index 2 from largest to smallest.
- ✓ Since there are more sites than the desired set size, then one accession could be chosen randomly from each site starting at the top ranked site until the desired set size is reached. Alternatively, this approach could be taken after one candidate accession is donated by each country represented in the candidate site list.

The climatic conditions layers were extracted from the GIS surfaces modeled from data collection sites as described by De Pauw (2008). The FIGS-LR subset has more accessions from Greece, Turkey, Ethiopia, and India (Supplementary Table S2).

Seedling screening

The seedling screening of GCP and FIGS_LR subsets was conducted under controlled conditions in the growth chamber with two pure isolates of LR (ISO-SAT and ISO-MRC). The single urediniospore was isolated from infected leaves collected from the experimental stations of Sidi Allal Tazi (ISO-SAT) and Marchouch (ISO-MRC) in 2017 and were multiplied on the susceptible barley cultivars (Bowman and Aglou) followed by collection and drying of urediniospores on silica gel and storage at -80 ºC until further use.

Barley plants were grown in sterilized peat moss (supplemented with 14–14–14 NPK) in plastic cones (14 cm long cones with 3.8 cm diameter) positioned in a 14 × 7-unit tray (Steuwe & Sons, Inc., OR, United States). For each barley accession, 4-5 seeds were planted per cone in two replications. Each tray contained 96-test genotypes along with resistant (Philadelphia) and susceptible (Bowman) checks. Plants were raised in the growth chamber (Snijder Scientific, Tilburg, the Netherlands) with a photoperiod of 16 h light/8 h dark at 20 ± 1ºC. Inoculation was carried out on 10 –12 days old seedlings when the first leaf was fully expanded. To prepare LR inoculum, urediospores were taken from the –80°C freezer and subjected to heat shock for 5 min at 40 °C.

For each tray, 15 mg of urediniospores were suspended in 10 ml of light mineral oil (Novec 7100, Sigma Aldrich), and this spore suspension was sprayed onto plants as a fine mist using an airbrush (Revell, Munchen, Germany). Inoculated plants were left to dry for 20 minutes at the room temperature and were placed in growth chamber in the dark for 24 hours at 18 °C with ~100% relative humidity. Then the plants were maintained in the growth

chamber with a light/ dark period of 16/8 h at 20 °C for symptoms development. The evaluation for LR disease was carried out 12-14 days post-inoculation based on infection types (ITs) according to the 0 to 4 scale developed by Stakman *et al.* (1962). The seedlings were classified either as immune (0), resistant (0; and 1), moderately resistant (2), moderately susceptible (3), or susceptible (4).

Field screening

The field phenotyping was performed at the INRA experimental station of Sidi Allal Tazi (34° 52' N, 6.32 W) during 2016-17 (SAT2017), 2017-18 (SAT2018), 2018-19 (SAT2019), at Marchouch (33° 56′10”N 6°69′21”W) during 2017-18 (MRC2018), and at Guich (33°58'59.7"N 6°51'41.6"W) during 2017-18 (GCH18). The tested accessions were sown as single rows of 1m with 0.5 m row spacing between adjacent accessions using an augmented block design. The seed mixture of susceptible cultivars Bowman and Aglou were sown as a rust spreader row at the border of each block to allow the uniform distribution of LR inoculum.

At SAT, the disease was established naturally, but at Guich station the disease was initiated using the artificial inoculation. About 1 gram of dried urediniospores were suspended in 200 ml of mineral oil and sprayed on the trial using an airbrush (Revell, Munchen, Germany). The establishment and spread of the disease were favored by covering the spreader rows with a plastic sheet overnight and by periodic sprinkler irrigation. The phenotypic evaluation for LR resistance was assessed for GCP and FIGS_LR subsets at growth stage 65-77 (Zadoks *et al.* 1974) using the modified Cobb scale (Peterson *et al.* 1948) which combined the LR severity (0 to 100%) and host response; 0 (Immune), No visible infection on plants; R (resistant), visible chlorosis or necrosis, no uredia are present; MR (moderately resistant), small uredia are present and surrounded by either chlorotic or necrotic areas; MS (moderately susceptible), medium sized uredia are present and possibly surrounded by chlorotic areas; S (susceptible), large uredia are present, generally with little or no chlorosis and no necrosis. The Coefficient of Infection (CI) was calculated by multiplying the infection response values (R = 0.2, MR = 0.4, MS = 0.8, S = 1) with the percent disease severity (0-100%) (Stubbs *et al.*1986), and the accessions were rated based on the average coefficient of infection (ACI) where values of 0-7, 8-16, 17-29, 30-50, and >50 were considered as resistant, moderately resistant, moderately susceptible, susceptible, and highly susceptible, respectively.

Data analysis

Comparing the reactions of GCP and FIGS subsets

The statistical analysis was performed using R software (R Core Team 2018). The statistical association between sub-setting approach and the reaction to LR was calculated using χ^2 test of independence with significance level ($\alpha = 0.05$) using the following equation:

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

The equation used for calculating expected values in a test of independence was as follows:

$$E_{ij} = \frac{\sum_{k=1}^{c} O_{ij} \sum_{k=1}^{r} O_{kj}}{N}$$

Where E_{ij} = expected value, $\sum_{k=1}^{c} O_{ij}$ is the sum of the ith column, $\sum_{k=1}^{r} O_{kj}$ is the sum of the kth row, N is the total number.

To find out the differences between FIGS and GCP subsets in terms of reaction to LR, the test of goodness of fit using χ^2 test at a significance level (α = 0.05) was used where GCP was simulated to a random sample. The expected values for the test of goodness of fit are calculated as follows:

$$E_i = np_i$$

Where E_i is the expected value, n is the total sample size, and p_i is the hypothesized proportion of observations in level *i*.

Both tests were performed using different groupings of reactions, all classes (I, R, MR, MS and S), three classes (I+R, MR+MS, S) and (I+ R+MR, MS, S), two classes (I+R+MR , MS+S) at the seedling stage, and all classes (R, MR, MS, S, HS), three classes (R, MR+MS, and S+HS), and two classes (R+MR+MS, S+HS) at the adult plant stage.

Modeling of the reaction to leaf rust disease

The second pathway of the focused identification of the germplasm strategy (FIGS) using machine learning was investigated using the available reactions of the accessions of FIGS_LR and GCP subsets to find a function that links adaptive traits, environments (and associated selection pressures) with genebank accessions. We used environmental data from WorldClim1 databases as predictors. The WorldClim is an open access database providing global climatic layers describing past climatic profiles of collection sites intended for spatial modeling or mapping. It includes averages of monthly minimum and maximum temperatures, precipitation and bioclimatic variables (Fick and Hijmans 2017).

The following machine learning algorithms were used: K-nearest neighbors KNN (Kotsiantis 2007), Support Vector Machine SVM (Hsu *et al.* 2010), Random Forest RF (Breiman, 2001), Neural networks NNET (Venables and Ripley 2002), and Bagged Carts BCART (Kołcz 2000). Each machine learning model was tuned to select the best tuning parameters using a training set (70% of the total set), and then the best model was selected between different machine learning models based on several metrics including accuracy, specificity, and Kappa. The modeling metrics were computed on the test set (30% of the total set). In this study, R language and caret library were used for machine learning analysis (Kuhn 2008). Models were tuned for parameter's optimization and trained on 70% of the data and tested with 10 cross validation folds and 100 replications. In addition, modeling was done for the two isolates for the seedling stage. For the APR, modeling was done for the entire multi-locations data sets and for each location separately.

Results

Seedling resistance

In the seedling test, successful artificial inoculation was carried out for the two isolates and diverse infection responses were recorded. The frequency distribution of infection response of GCP and FIGS_LR accessions at the seedling stage has been presented in Figure 1. The uniformity of the disease development was assessed through the high susceptibility of the check Bowman over the two replications. The highest number of accessions of both FIGS_LR (65.88 % for ISO-MRC, and 51.16 % for ISO-SAT), and GCP (62.87 and 48.4 % for ISO-MRC and ISO-SAT isolates, respectively) subsets showed MR reaction, and only few accessions showed R reaction with one accession of GCP having immune reaction (Figures 1a and 1b). More resistant accessions were noted when tested to ISO-MRC isolate with 5.88 % for FIGS_LR and 22.75 % for GCP compared with ISO-SAT isolate that showed 10.47 % and 13.3 % resistant accessions for FIGS_LR and GCP, respectively. Some accessions showed different disease reaction for the two isolates. For example, the accessions IG 143872, IG 143862, IG 143864, IG 143872, IG 143890, IG 143978, IG 143984, IG 144006, IG 144090, IG 144029, IG 144077, IG 143871, and IG 144012 were resistant to ISO-MRC isolate but not to ISO-AT isolate, whereas accessions IG 143963, IG 18725, IG 19525 were resistant to ISO-AT isolate but not to ISO-MRC isolate. Ten barley accessions in GCP (IG 143876, IG 143886, IG 143906, IG 143929, IG 143998, IG 143999, IG 144014, IG 144064, IG 144076, IG 144108) and one in FIGS_LR (IG 18957) were resistant to both isolates. Most of the resistant accessions originated from USA, Turkey, Greece, and Morocco.

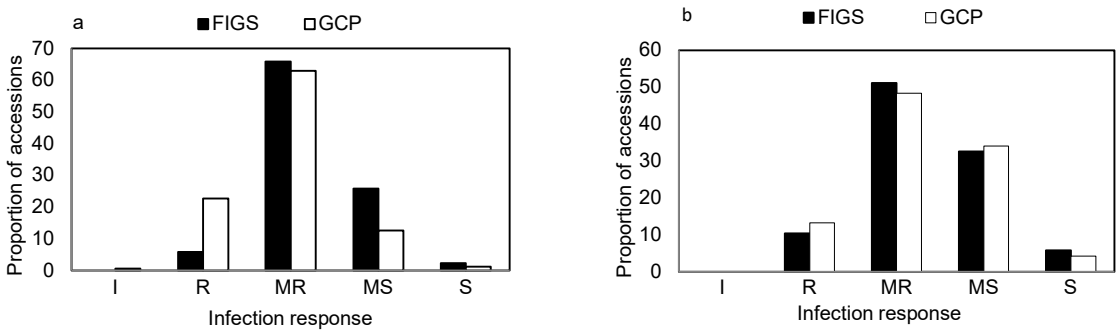

Fig. 1 Frequency distributions of LR disease reactions of barley FIGS_LR and GCP subsets evaluated at the seedling stage against ISO-MRC (a) and ISO-SAT isolates (b). Here I = immune, R = resistant, MR = moderately resistant, MS = moderately susceptible, and S = susceptible.

Adult Plant Resistance (APR)

Under field conditions, good natural LR infection was obtained at the Sidi Allal Tazi during three cropping seasons, and good artificial infection was established at Guich in 2018. However, late and light artificial infection at Marchouch during 2017 and 2018 seasons did not allow disease severity assessments. The uniformity of the disease development was assessed through the high susceptibility of the checks, Bowman and Aglou, at the adult stage in Sidi Allal Tazi and Guich sites. The good development of the disease allowed efficient screening of the germplasm

at the adult plant stage, as shown by wide range of reactions observed (Figure 2). The average of coefficient of infection (ACI) values across the environments ranged from 0 to 85 with several accessions showing contrasting reactions in different environments. Accessions of FIGS_LR and GCP subsets showed different distributions of the reaction classes with near normal distribution for SAT2017 and SAT2019, and positive skewness with high percentage of R accessions (ranged from 55.8 to 68.3%) for SAT2018 and GCH2018 (Figures 2b and 2c). While at SAT2017 and SAT2019, this rate ranged from 5.59 to 19.64% (Figures. 2a and 2d) for both subsets. When considering MR reaction, additional 8.94 to 30.36% accessions were found during SAT2017 and SAT2019 cropping seasons. The highest numbers of susceptible accessions were observed at SAT2017 with 34.64 and 15.29% showing susceptibility, and 22.91% and 15.29% were highly susceptible (HS) for GCP and FIGS_LR, respectively. Only three accessions in FIGS_LR (IG 28636, IG 28647 and IG 33039), and four accessions in GCP subset (IG 143945, IG 144000, IG 144064, IG 144105) were found to be resistant across all environments (Table 1).

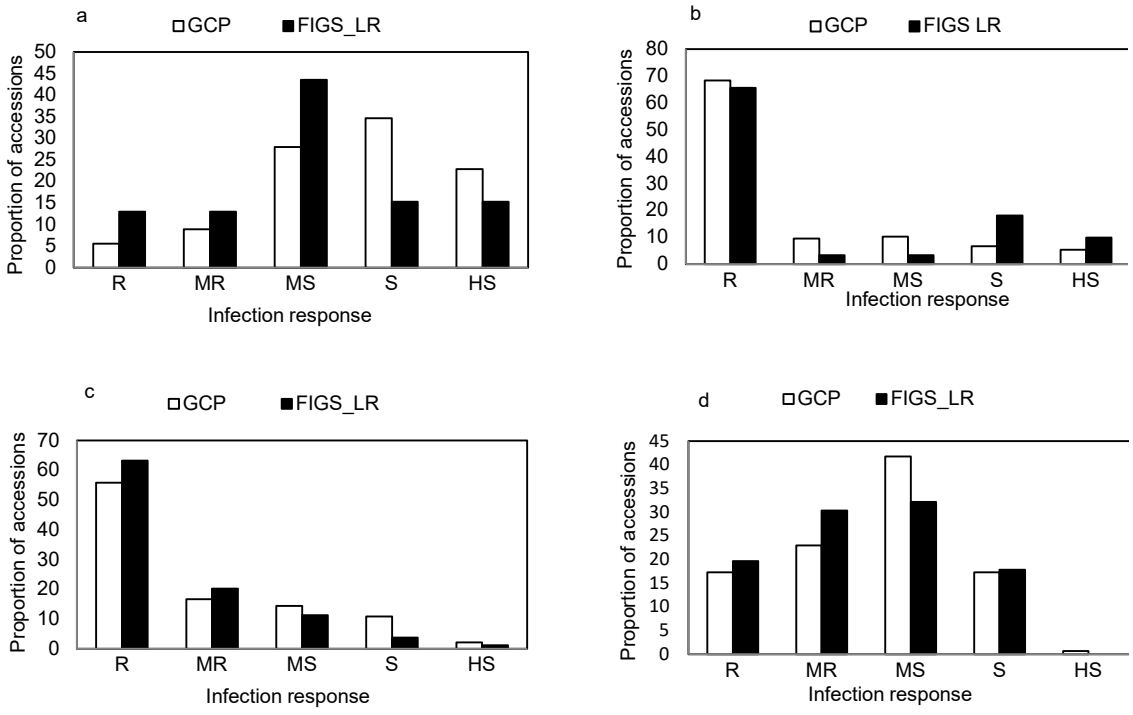

Fig. 2 Frequency distributions of the adult plant resistance (APR) of FIGS_LR and GCP subsets to leaf rust under field conditions at Sidi Allal Tazi during (a) 2017 (SAT2017), (b) 2018 (SAT2018), (d) 2019 (SAT2019) cropping seasons, and at Guich during 2018 (GCH2018 (c)). Here R = resistant, MR= moderately resistant, MS = moderately susceptible, and S = susceptible.

Table 1 Average of coefficient of infection (ACI) values of most resistant accessions at the adult plant stage across environments, and their infection responses (IR) at the seedling stage to LR isolates (ISO-MRC, ISO-SAT).

Accession	Subset	SAT2017	SAT2018	GCH2018	SAT2019	ISO-MRC	ISO-SAT
IG 143945	GCP	8	6.2	2.5	6	2	2
IG 144000	GCP	12	2	6.8	4	2	2
IG 144064	GCP	2	0.8	0.4	6	1	0;
IG 144105	GCP	4.2	0.4	1	12	2	3
IG 28636	FIGS	6	1.1	4.95	5	2	2
IG 28647	FIGS	6.1	4.1	0.7	8	2	2+
IG 33039	FIGS	4.2	0.4	6.7	0.2	2	2

Comparison of the reaction of FIGS_LR and GCP to leaf rust

For ISO-MRC isolate at the seedling stage, GCP subset showed higher percentage of resistant accessions (22.75 %) than FIGS_LR (6%). Similarly, 13.3% of GCP accessions and 10.5% of FIGS_LR accessions were resistant to ISO-SAT isolate. While approximately similar percentages were found for MR and S reaction classes for both isolates (Figures 1a and 1b). At the adult plant stage, FIGS_LR had slightly higher percentages of accessions with R and MR reactions except at Sidi Allal Tazi during 2017-18 (SAT2018) season (Figure. 1b). Under severe epidemic of LR as observed at Sidi Allal Tazi during 2016-17 season (SAT2017), FIGS_LR showed less accessions having S and HS reactions compared to GCP subset (Figures. 1a and 1c).

Applying the chi-square tests of independence to both subsets, the results showed that there is a significant relationship (P-value=0.04) between the response to LR disease and the sub-setting approach for ISO-MRC isolate, but not for ISO-SAT, when considering all classes of reaction. For ISO-MRC, GCP included one accession with immune reaction and 23% of the accessions being resistant, while FIGS_LR subset showed only 6% of the accessions being resistant (Table 2, Figure 1a). When screened with ISO-SAT isolate, both FIGS_LR and GCP showed that the infection response to LR was not dependent on the sub-setting (p=0.85), displaying same distribution patterns of the reaction. No accession was found immune in both subsets while 9 (10.47%) and 25 (13.3%) accessions were resistant, and 44 (51.16%) and 91 (48.4%) were MR for FIGS_LR and GCP subsets, respectively (Table 2, Figure 1b).

The test of goodness of fit showed that FIGS_LR subset yielded higher percentages of accessions with R, and MR reactions than GCP in case of ISO-MRC isolate, but no significant differences were observed between the two subsets when tested with ISO-SAT isolate.

Table 2 Number of accessions per reaction type of barley FIGS_LR and GCP subsets evaluated at the seedling stage using two leaf rust isolates (ISO-MRC and ISO-SAT) with $\chi 2$ (P-values) for the tests of independence and goodness of fit.

	ISO-MRC		ISO-SAT	
	FIGS_LR	GCP	FIGS_LR	GCP
I	0	1	0	0
R	5	38	9	25
MR	56	105	44	91
MS	22	21	28	64
S	2	2	5	8
*$\chi 2$ (P)	16 .31 (0.04)		0.81 (0.85)	
**$\chi 2$ (P)	24.18 (0.0001)		1.2 (0.75)	

*$\chi 2$ test of independence, and **$\chi 2$ goodness of fit test.

Here I: immune; R: resistant; MR: moderately resistant; MS: moderately susceptible; S: susceptible.

Except the grouping of two classes (I+R+MR, MS+S) for the goodness of fit test, the tests of independence and goodness of fit were significant for ISO-MRC, but not for ISO-SAT for different groupings of the reactions (Table 3).

Table 3 $\chi 2$ (P-value) for the tests of independence and goodness of fit for different groupings of the reaction to two leaf rust isolates (ISO-MRC and ISO-SAT) at the seedling stage for FIGS-LR and GCP subsets.

Groups of reaction	Test of independence		Test of goodness of fit	
	ISO-MRC	ISO-SAT	ISO-MRC	ISO-SAT
I+R, MR+MS, S	14.65 (0.001)	0.70 (0.71)	15.13 (0.001)	1.03 (0.60)
I+ R+MR, MS, S	11.67 (0.003)	0.34 (0.84)	9.54 (0.008)	0.55 (0.76)
I+R+MR, MS+S	10.10 (0.001)	0.0001 (0.99)	6.61 (0.10)	0.0002 (0.99)

Here I: immune; R: resistant; MR: moderately resistant; MS: moderately susceptible; S: susceptible.

For APR, the reaction to LR was dependent on subsets for Sidi Allal Tazi during 2017 (SAT2017) and 2018 (SAT2018) with respective $\chi 2$ (P-values) of 0.001 and 0.02, respectively. But this dependence was not found in case of GCH 2018 (GCH18) and Sidi Allal Tazi 2019 (SAT2019) (Table 4). The tests of goodness of fit showed that FIGS-LR outperformed GCP subset at Sidi Allal Tazi with higher percentages of accessions with R and MR reactions under heavy infection in 2017, but the opposite was observed during 2018 season at the same site. For GCH2018 and SAT2019 environments, no significant differences were observed between the two subsets.

Table 4 Number of barley accessions per leaf rust reaction type of FIGS-LR and GCP subsets evaluated at the adult plant stage at Sidi Allal Tazi (SAT) and Guich (GCH) stations with $\chi 2$ (P-values) for the tests of independence and goodness of fit.

	SAT2017		SAT2018		GCH2018		SAT2019	
	FIGS_LR	GCP	FIGS_LR	GCP	FIGS_LR	GCP	FIGS_LR	GCP
R	11	10	40	114	50	77	11	24
MR	11	16	2	16	16	23	17	32
MS	37	50	2	17	9	20	18	58
S	13	62	11	11	3	15	10	24
HS	13	41	6	9	1	3	0	1
*$\chi 2$ (P)	18,3 (0,001)		12,26 (0,02)		4,46 (0,35)		2,33 (0,67)	
**$\chi 2$ (P)	28,49 (9,94E^{-06})		19,82 (0,001)		5,86 (0,21)		3,14 (0,53)	

*$\chi 2$ test of independence and **$\chi 2$ goodness of fit test

Here R: resistant; MR: moderately resistant; MS: moderately susceptible; S: susceptible; HS: highly susceptible.

When different groupings of the reactions were performed, the significance of the two tests were highly significant for SAT2017 and SAT2018, but not for GCH2018 and SAT2019, except for the test of goodness of fit for GCH2018 in case of the grouping (R+MR+MS; S+HS) with P-value of 0.04 (Table 5).

Table 5 $\chi 2$ (P-value) for the tests of independence and goodness of fit of the leaf rust reactions at the adult plant stage of FIGS-LR and GCP subsets of barley at Sidi Allal Tazi (SAT2017, SAT2018 and SAT2019) and Guich (GCH2018).

Group of reactions	SAT2017	SAT2018	SAT2019	GCH2018
	Test of independence			
R, MR+MS, S+HS	17.61 (0.0001)	11.8 (0.003)	0.16 (0.92)	3.64 (0.162)
R+MR+MS, S+HS	16.76 (4.25E-05)	8.30 (0.004)	0.0004 (0.98)	3.51 (0.061)
Group of reactions	Test of goodness to fit			
R, MR+MS, S+HS	27.82 (9.11E-07)	18.31 (0.0001)	0.23 (0.89)	4.66 (0.097)
R+MR+MS, S+HS	25.28 (4.97E-07)	14.62 (0.0001)	0.001 (0.98)	4.44 (0.04)

Here R: resistant; MR: moderately resistant; MS: moderately susceptible; S: sensible; HS: highly susceptible

Predictive modeling of the reaction to leaf rust

At the seedling stage, the tested machine learning models did not perform similarly for the two LR isolates. For ISO-MRC isolate, all models yielded a significant medium to very high accuracy. The maximum accuracy (0.94) was reached using the BCART model and was then chosen as the best model. The remaining modeling parameters showed the strong mathematical relationship between the reaction to ISO-MRC and the environmental characteristics (Table 6). However, the modeling pattern was opposite for ISO-SAT isolate where all the models were not significantly accurate (Table 7), since the accuracy was similar to the "No Information Rate" and hence

demonstrating that the models were as good as the naïve model. It is noticeable that the specificity was much lower than sensitivity for all tested models.

Table 6 Performance measures of different machine learning models for resistance of barley accessions to leaf rust isolate ISO-MRC at the seedling stage.

Performance Measures	k-NN	SVM	RF	NNET	BCART
Accuracy	0.716	0.686	0.961	0.912	0.941
95% CI	(0.62-0.80)	(0.59-0.77)	(0.90- 0.99)	(0.84- 0.96)	(0.88- 0.98)
No Information Rate	0.5	0.5	0.5	0.5	0.5
P-Value [Acc > NIR]	7.75E-06	0.000107	8.73E-25	4.99E-19	2.83E-22
Kappa	0.431	0.373	0.922	0.824	0.882
Sensitivity	0.549	0.529	0.922	0.824	0.882
Specificity	0.882	0.843	1	1	1

Here k-NN: K-nearest neighbors KNN; SVM: Support Vector Machine; RF: Random Forest; BCART: Bagged Carts

Table 7 Performance measures of different machine learning models for resistance of barley accessions to leaf rust isolate ISO-SAT at the seedling stage.

Performance Measures	k-NN	SVM	RF	NNET	BCART
Accuracy	0.722	0.556	0.597	0.542	0.653
95% CI	(0.60-0.82)	(0.43-0.67)	(0.47-0.71)	(0.42-0.66)	(0.53-0.76)
No Information Rate	0.639	0.639	0.639	0.639	0.639
P-Value [Acc > NIR]	0.086782	0.943047	0.805548	0.965561	0.455707
Kappa	0.355	0.0368	0.089	-0.0189	0.228
Sensitivity	0.87	0.652	0.739	0.674	0.761
Specificity	0.462	0.385	0.346	0.308	0.462

Here k-NN: K-nearest neighbors KNN; SVM: Support Vector Machine; RF: Random Forest; BCART: Bagged Carts

For the APR, no model performed significantly for the two locations (Table 8). Accuracy was higher for all models, however, the unbalanced data due to the higher number of resistant genotypes make the model not performing better than the naïve model because of the low values of specificity and high value of "No Information Rate". Among the tested models, RF was the best model for all locations.

Table 8 Performance measures of the best machine learning model (Random Forest) for the prediction of leaf rust adult plant resistance of barley accessions.

Model	(SAT+ GCH)	SAT	GCH
	RF	RF	RF
Accuracy	0.81	0.81	0.84
95% CI	(0.741, 0.864)	(0.691, 0.903)	(0.719, 0.918)
No Information Rate	0.80	0.75	0.85
P-Value [Acc > NIR]	0.47	0.15	0.72
Kappa	0.31	0.42	0.20
Sensitivity	0.92	0.96	0.94
Specificity	0.35	0.40	0.22

Here SAT: Allal Tazi; GCH: Guich; RF: Random Forest

Discussion

LR occurs annually with high incidence in the Northern regions of Morocco, and the Sidi Allal Tazi has been used as the disease hotspot for the barley germplasm screening. Over three years, most of the barley varieties and advanced breeding lines showed high susceptibility to the population of *P. hordei* at this site. The development of high yielding varieties with adequate levels of resistance is the key to an integrated management of LR and is compatible with the general consideration of barley as a low input crop by most farmers. Recent studies on the screening of international germplasm collections reported that the effective sources of LR resistance available are limited, highlighting the needs for additional new sources of resistance (Golegaonkar *et al*. 2009; Derevnina *et al*. 2013; Sandhu *et al*. 2014, Singh *et al*. 2015). Genetic resources conserved *ex situ* in the genebanks are important sources of breeders 'sought traits including the resistance to major diseases, but need efficient mining approaches. In this study, both FIGS_LR and GCP subsets have yielded sources of resistance to LR, but only few accessions showed resistant reaction at the seedling and/or adult plant stages under heavy infection and high virulence of isolates and populations of the pathogen as is the case of SAT2017 and SAT2019.

The results indicated the changes in the reaction of accessions to different isolates at the seedling stage and at different locations and in different years at the adult plant stage. There were few accessions (IG 143886, IG 143906, IG 143999, IG 144108, IG 143963, IG 143990, IG 144071, IG 144012, IG 144019, IG 144085, and IG 144148) which showed resistance at the seedling stage but were susceptible at the adult plant stage under field conditions.

The seedling resistance is usually characterized by hypersensitivity and is governed by single major genes (race-specific). To date 23 hypersensitive seedling resistance genes have been identified (*Rph1* - *Rph19*, *Rph21*, *Rph22*, *Rph25* and *Rph 26*) (Park *et al.* 2015, Kavanagh *et al.* 2017; Yu *et al*. 2018). Such genes can be easily overcome by new LR races because of their excessive utilization over large areas which exert selection pressure on the

pathogen population which lead to the emergence of new races, and eventual breakdown the effectiveness of resistance genes. Several studies reported that the frequency of virulence for the resistance gene *Rph4* was increased in Queensland during the 1980s, possibly due to the widespread use of the cultivar Grimmett, which possesses *Rph4* (Cotterill *et al.* 1995; Burdon *et al.* 2014; Li *et al.* 2014; Niks *et al.* 2015). Woldeab *et al.* (2006) reported the detection of eleven new pathotypes since 1992-2001 in Australia and seven new pathotypes in Ethiopia in 2003 and 2004. Virulence has been detected for most known seedling *Rph* genes in various barley growing regions throughout the world. In Australia, only *Rph3*, *Rph7*, *Rph11*, *Rph14*, *Rph15*, and *Rph18* of the characterized major genes were still effective to prevailing pathotypes (Cotterill *et al*. 1995; Park 2003). However, pathotypes virulent to *Rph3* were detected in New Zealand (Cromey and Viljanen-Rollinson 1995), and the virulence for *Rph7* has been identified in Israel (Golan *et al.*1978), Morocco (Parlevliet *et al*. 1981), and North America (Steffenson *et al*. 1993). Virulence for *Rph11* and *Rph14* has also been found frequently in many parts of the world (Fetch *et al.* 1998), and virulence to *Rph15* was reported by Sun *et al.* (Sun 2007). Therefore, an accession with LR resistance at the seedling stage alone might not provide durable and effective resistance (Singh 1992; Park 2008; Singh *et al.* 2015).

Only few accessions; IG 143945, IG 144000, IG 144140, IG 144064 from GCP subset, and accessions IG 28613, IG 28636, IG 33039 from FIGS_LR subset showed resistant (R) to moderately resistant (MR) response at the adult plant stages across all environments. Of the 21 Bowman differential lines tested at the adult plant stage at Sidi Allal Tazi, only one-line displayed moderately resistant reaction to LR field population (unpublished data). Most probably, the LR resistant accessions identified in this study may possess either new *R* genes or allelic variant of existing *R* genes or a combination of both. A high-density genotyping and genome wide association studies seem to be logical step to dissect the resistance diversity. These putative *R* genes could be either pyramided or used sequentially to ensure a better *R* gene deployment strategy.

Since there are several accessions at the adult plant stage with MR and MS reactions or with slow progression of the disease based on the area under the disease progress curve (data not presented) under heavy rust epidemics, partial resistance and slow rusting mechanism could be considered to ensure a race non-specific and a more durable resistance. Several studies have promoted partial and non-race specific resistance in case of rusts and powdery mildew in barley and wheat as this type of resistance is available in some commercial varieties (Parlevliet and Kuiper 1977; Andres and Wilcoxon 1986; Niks *et al.* 2000; Stuthman *et al.* 2007). Several APR genes were well characterized and deployed in wheat to control rust diseases (Park and McIntosh 1994). In barley, three genes governing APR to LR have been identified and used (*Rph20*, *Rph23*, and *Rph24*) (Hickey *et al.* 2011; Singh *et al.* 2015; Ziems *et al*. 2017). Even if there are no reports of virulence for *Rph20*, *Rph23* or *Rph24*, identifying new APR resistance genes for LR are essential for diversifying resistance and to promote gene pyramiding to increase resistance levels. Marker assisted selection (MAS) provides an opportunity to breeders to pyramid the APR genes in barley.

The Focused Identification of Germplasm Strategy (FIGS) has been promoted by ICARDA genebank as an innovative approach for efficient mining the genebank holdings to better respond to seed requests for specific traits sought by breeders. This approach has shown its efficiency in identifying novel sources of resistance to powdery

mildew, yellow and stem rusts, Sunn pest, and Russian wheat aphid in wheat (Bhullar *et al.* 2009; El Bouhssini *et al.* 2009, 2011; Bari *et al*. 2012, 2014), and to net blotch of barley (Endresen et al. 2011). This study included the first attempt to compare FIGS with another subset, the Reference set of the Generation Challenge Program (GCP) selected from the global barley core collection based on diversity using SSR markers. The tests of independence showed that the dependence of LR reaction on sub-setting approach varied among the isolates at the seedling stage, and between environments at the adult plant stage. It appears that the dependence is mainly found for ISO-MRC isolate and under heavy infection level experienced at Sidi Allal Tazi. Similarly, the tests of goodness of fit showed that the differences between FIGS-LR and GCP for percentages of accessions with different reactions varied with the isolate at the seedling stage, and on sites and years for the adult plant stage. FIGS sub-setting using filtering approach has allowed to identify higher percentages of accessions with R and MR reactions compared to GCP subset in case of field tests (except SAT2018). Modeling outcomes using machine learning approach were dependent on the isolates, and the environments, and associated pathogen field populations. The reduced sample size as well as the non-balance between the two classes (Resistant and Susceptible) could explain the limitation of the machine learning models. The results showed the need for further fine tuning of FIGS approach to consider the diversity of virulence of the pathogen populations using larger subsets. Since the GCP reference set included a larger number of accessions with R and MR reactions at both the seedling and the adult plant stages. It will be interesting to compare both sub-setting methods in yielding new different effective genes. This can be investigated using molecular markers or by screening the identified sources of resistance to a larger number of isolates with different virulence spectrums.

Conclusion

This current study suggests that the trait mining approach can be an efficient alternative to the core collection method. The resistant and moderately resistant accessions at the seedling and at the adult plant stages in this study are valuable resources of *P. hordei* resistance and can lead towards effective and durable resistance against *P. hordei* when combined with appropriate gene deployment strategies. The evaluation of larger subsamples in different environments, and against different pathotypes will allow the fine tuning of FIGS sub-setting approach using machine leaning.

Supplemental Table S1. List of 188 accessions from the reference set constructed within the Generation Challenge Program (GCP) used in the study and relevant information such as IG #, Origin, and collection site.

Supplemental Table S2. List of 86 barley accessions selected using filtering approach of the Focused Identification of Germplasm Strategy (FIGS_LR) and relevant information such as IG #, Origin, and collection site.

Declarations

Acknowledgements: The authors would like to thank Dr. Kenneth Street for his help in developing FIGS subset and Mr. Amer El-Omrani, the technician at Sidi Allal Tazi experimental station for his help with field experiments.

Funding: This study is supported financially through GIZ-attributed funding to ICARDA genebank and the CAIGE-GRDC-ICA00010 project.

Data availability statement

The data that support the findings of this study are available in the ICARDA genebank database and can be obtained upon request from the corresponding author.

Compliance with ethical standards

This work does not contain any studies with human participants or animals performed by any of the authors

Conflict of interest

The writers declare that they have no known conflicting financial interests or personal relations that may have had an impact on the work presented in this article.

References

Andres MW, Roy DWilcoxson (1986) Selection of Barley for Slow Rusting Resistance to Leaf Rust in Epidemics of Different Severity. Crop Sci 26: 511-514. https://doi.org/10.2135/cropsci1986.0011183X002600030015x

Arnst BJ, Martens JW, Wright GM, Burnett PA, Sanderson FR (1979) Incidence, importance and virulence of *Puccinia hordei* on barley in New Zealand. *Ann Appl Biol* 92:185–90

Bari A, Street K, Mackay M, Endresen DTF, De Pauw E (2012) Focused identification of germplasm strategy (FIGS) detects wheat stem rust resistance linked to environmental variables. Genet Resour Crop Evol 59: 1465–1481.

Bari A, Amri A, Street K, Mackay M, De Pauw E, Sanders R, Nazari K, Humeid B, Konopka J, Alo F (2014) Predicting resistance to stripe (yellow) rust (Puccinia striiformis) in wheat genetic resources using focused identification of germplasm strategy. The Journal of Agricultural Science 152(06):906–916. DOI: 10.1017/S0021859613000543

Bhullar NK, Street K, Mackay M, Yahiaoui N, Keller B (2009) Unlocking wheat genetic resources for the molecular identification of previously undescribed functional alleles at the Pm3 resistance locus. Proc Natl Acad Sci USA 106: 9519–9524.

Breiman L (2001) Random forests. Machine Learning 45(1):5–32

Brunner S, Keller B, Feuillet C (2000) Molecular mapping of the *Rph7.g* leaf rust resistance gene in barley (*Hordeum vulgare* L.). Theor Appl Genet 101:783–788

Burdon JJ, BarrettLG, RebetzkeG, Thrall PH (2014) Guiding deployment of resistance in cereals using evolutionary principles. Evol Appl 7:609–624. 10.1111/eva.12175

Chicaiza O, Franckowiak JD, and Steffenson BJ (1996) New sources of resistance to leaf rust in barley. In: Slinkard AE, Scoles GJ, Rossnagel BG (eds) Proc Fifth Int Oat Conf & Seventh Int Barley Genet Symp. Saskatoon, pp 706-708.

CGIAR Generation Challenge Programme (2012) Paper No 3: Genetic stocks). Position Papers: GCP's research component. Texcoco, Mexico. 22 pp.

Clifford BC (1985) Barley leaf rust. In: Roelfs AP, Bushnell WR (eds) The cereal Rusts. Diseases, Distribution, Epidemiology and Control, Vol II. Harcourt Brace Jovanovich, Publishers Orlando, Florida 32887, pp 173-205

Cotterill PJ, Rees RG, Platz GJ, Dill-Macky R (1992) Effects of leaf rust on selected Australian barleys. Aus J Exp Agric 32: 747-751

Cotterill PJ, Park RF, ReesRG (1995) Pathogenic specialization of *Puccinia hordei* Otth. in Australia, 1966-1990. Aus J Agric Res 46: 127-134

Cromey MG, Viljanen-Rollinson SLH (1995) Virulence of *Puccinia hordei* on barley in New Zealand from 1990 to 1993. N Z J Crop Hortic Sci 23:115–19

Derevnina L, Singh D, Park RF (2013) Identification and characterization of seedling and adult plant resistance to *Puccinia hordei* in Chinese barley germplasm. Plant Breeding 132(6): 571–579. doi:10.1111/pbr.12082

De Pauw, Oweis T, Youssef J (2008) Integrating expert knowledge in GIS to locate biophysical potential for water harvesting: Methodology and a case study for Syria. ICARDA, Aleppo, Syria pp 59

596 Dwivedi SL, Stalker HT, Blair MW, Bertioli D, Upadhyaya HD, Nielen S et al. (2008) Enhancing crop gene pool
597 with beneficial traits using wild relatives. Plant Breed 30:179–230. 10.1002/9780470380130.
598 El Bouhssini M, Street K, Joubi A, Ibrahim Z, Rihawi F (2009) Sources of wheat resistance to sunn pest,
599 *Eurygaster integriceps* Puton, in Syria. Genet Resour Crop Evol 56: 1065–1069
600 El Bouhssini M, Street K, Amri A, Mackay M, Ogbonnaya FC et al (2011) Sources of resistance in bread wheat
601 to Russian wheat aphid (*Diuraphis noxia*) in Syria identified using the Focused Identification of Germplasm
602 Strategy (FIGS). Plant Breed 130: 96–97
603 Elmansour H, Singh D, Dracatos PM, Park RF (2017). Identification and characterization of seedling and adult
604 plant resistance to *Puccinia hordei* in selected African barley germplasm. Springer Science+Business Media
605 Dordrecht. DOI 10.1007/s10681-017-1902-8
606 Endresen DTF, Street K, Mackay M, Bari A, De Pauw E (2011) Predictive association between biotic stress traits
607 and ecogeographic data for wheat and barley landraces. Crop Sci 51: 2036–2055
608 Endresen DTF, Street K, Mackay M, Bari A, Amri A, De Pauw E, Nazari K, Yahyaoui A, (2012) Sources of
609 resistance to stem rust (ug99) in bread wheat and durum wheat identified using focused identification of
610 germplasm strategy (FIGS). Crop Science 52: 764–773
611 FAO. FAOStat, http://www.fao.org/faostat/en/#rankings/countries_by_commodity (2017)
612 Fetch TG, Steffenson BJ, Jin Y (1998) Worldwide virulence of *Puccinia hordei* on barley. Phytopathology 88:S28
613 Fick SE and Hijmans RJ (2017) WorldClim 2: new 1km spatial resolution climate surfaces for global land areas.
614 International Journal of Climatology 37 (12): 4302-4315
615 Figueroa M, Upadhyaya NM, Sperschneider J, Park RF, Szabo LJ, Steffenson B, Dodds PN (2016) Changing the
616 Game: Using Integrative Genomics to Probe Virulence Mechanisms of the Stem Rust Pathogen *Puccinia*
617 *graminis* f. sp. *tritici*. Frontiers in Plant Science 1–10. http://doi.org/10.3389/fpls.2016.00205
618 Frankel OH, Brown AHD (1984) Current plant genetic resources- a critical appraisal. In: Chopra VL, Joshi BC,
619 Sharma RP, and Bansal HC (eds) Genetics: New Frontiers. (Oxford & IBH Publ Co, New Delhi, India, pp 1-
620 13
621 Global Crop Diversity Trust (2008) Global strategy for the ex situ conservation and use of barley germplasm
622 [online]. Available from http://www.croptrust.org/documents /web/ BarleyStrategy_FINAL_27Oct08.pdf
623 Golan T, Anikster Y, Moseman JG, Wahl I (1978) A new virulent strain of *Puccinia hordei*. Euphytica 27:185–
624 89
625 Golegaonkar, PG, Singh D, Park RF (2009) Evaluation of seedling and adult plant resistance to *Puccinia hordei*
626 in barley. Euphytica 166: 183—197
627 Gollin D, Smale M, Skovmand B (2000) Searching an ex situ collection of wheat genetic resources. Am J Agric
628 Econ 82: 812–827
629 Gollin, D., M. Smale, and Skovmand B (2000) Searching an ex situ collection of wheat genetic resources.
630 American Journal of Agricultural Economics 82(4):812–827. https://doi.org/10.1111/0002-9092.00083
631 Griffey CA, Das MK, Baldwin RE, Waldenmaier CM (1994) Yield losses in winter barley resulting from a new
632 race of *Puccinia hordei* in North America. Plant Dis 78:256–260
633 Hajjar R, and Hodgkin T (2007) The use of wild relatives in crop improvement: A survey of developments over
634 the last 20 years. Euphytica 156:1–13. doi:10.1007/s10681-007-9363-0

635 Hickey LT, Lawson W, Platz GJ, Dieters M, German S, et al. 2011. Mapping *Rph20*: a gene conferring adult
636 plant resistance to *Puccinia hordei* in barley. Theor Appl Genet 123:44–68
637 Hintum Th van, Menting F (2003) Diversity in ex situ genebank collections of barley. In: Bothmer R von, Hintum
638 Th van, Knüpffer H and Sato K (eds) diversity in Barley (*Hordeum vulgare*). Elsevier Science BV, Amsterdam,
639 The Netherlandspp 247-257
640 Hsu CW, Chang CC, Lin CJ (2010) A Practical Guide to Support Vector Classification, Department of Computer
641 Science, National Taiwan University, Taipei, Taiwan
642 ICARDA (2013) A new approach to mining Agricultural gene banks – to Speed the pace of research Innovation
643 for food security. "FIGS" – the Focused Identification of Germplasm Strategy. Research in Action 3.
644 International Center for Agriculture Research in Dry Areas, Beirut, Lebanon.
645 ICARDA (2015) ICARDA Annual Report 2014. International Center for Agricultural Research in the Dry Areas,
646 Beirut, Lebanon. 56 pp. ISSN: 92-9127-290-6
647 Kavanagh PJ, Singh D, Bansal UK, and Park RF (2017) Inheritance and characterization of the new and rare gene
648 *Rph25* conferring seedling resistance in *Hordeum vulgare* against *Puccinia hordei*. Plant Breed 136:908–912
649 Khazaei H, Street K, Bari A, Mackay M, Stoddard FL (2013) The FIGS (Focused Identification of Germplasm
650 Strategy) Approach Identifies Traits Related to Drought Adaptation in Vicia faba Genetic Resources. PLoS
651 ONE 8(5): e63107. https://doi.org/10.1371/journal.pone.0063107
652 Knüpffer H, Hintum Th van (1995) The barley core collection: an international effort. In: (T. Hodgkin T, Brown
653 AHD, Hintum Th van and Morales EAV (eds) Core Collections of Plant Genetic Resources John Wiley and
654 Sons, Chichester, UK. pp. 171-178
655 Kołcz A (2000) N-tuple Network, CART, and Bagging, in Neural Computation,
656 Kotsiantis SB (2007) Supervised machine learning: a review of classification techniques, Informatica
657 Kuhn M (2008) Building Predictive Models in R Using the caret Package, Journal of Statistical Software 28: 1–
658 26
659 Li LG, Cai L, Zhang XX, Zhang T (2014) Potentially novel copper resistance genes in copper-enriched activated
660 sludge revealed by metagenomic analysis. Appl Microbiol Biotechnol 98: 10255–10266.
661 Mackay M (1990) Strategic planning for effective evaluation of plant germplasm. In: Srivastava JP, Damania AB
662 (eds) Wheat genetic resources: meeting diverse needs. John Wiley & Sons, Chichester, pp 21–25
663 Mackay M (1995) One core collection or many? In: Hodgkin T, Brown AHD, van Hintum TJL, Morales EAV
664 (eds) Core collections of plant genetic resources. John Wiley & Sons Ltd, Chichester, pp 199–210
665 Mackay M, Street K (2004) Focused identification of germplasm strategy – FIGS. In: Black CK, Panozzo JF,
666 Rebetzke GJ (eds) Proceedings of the 54th Australian Cereal Chemistry Conference and the 11th Wheat
667 Breeders' Assembly, pp 138–141. Royal Australian Chemical Institute (RACI), MelbourneMcIntosh R (1988)
668 The Role of Specific Genes in Breeding for Durable Stem Rust Resistance in Wheat and Triticale. In:
669 Simmonds S, Rajaram NW (eds)Breeding strategies for resistance to rust of wheat pp 1–9. Mexico, DF:
670 CIMMYT.
671 Niks RE, Walther U, Jaiser H, Martínez F, Rubiales D, Andersen O, Flath K, Gymer P, Heinrichs F, Jonsson R,
672 Kuntze L, Rasmussen M, Richter E (2000) Resistance against barley leaf rust (*Puccinia hordei*) in West-
673 European spring barley germplasm. Ann appL Bid 92: 185-190

Niks RE, Qi XQ, Marcel TC (2015) Quantitative resistance to biotrophic filamentous plant pathogens: concepts, misconceptions and mechanisms. in: vanAlfen NK (ed) Annual Review of Phytopathology pp 445–470

Park RF, McIntosh RA, (1994) Adult plant resistances to *Puccinia recondita* f. sp. *tritici* in wheat. N Z J Crop Hortic Sci 22:151–158

Park RF (2003) Pathogenic specialization and pathotype distribution of *Puccinia hordei* in Australia, 1992 to 2001. Plant Dis 87:1311–16

Park RF, (2008) Breeding cereals for rust resistance in Australia. Plant Pathology 57: 591–602. 10.1111/j.1365-3059.2008.01836.x

Park RF, Golegaonkar P, Derevnina L, Sandhu K, Karaoglu H, Elmansour H, Dracatos P, Singh D (2015) Leaf Rust of Cultivated Barley: Pathology and Control. Annual Review of Phytopathology 53: 565-589

Parlevliet JE, Kuiper HJ (1977) Partial resistance of barley to leaf rust, *Puccinia hordei*. IV. Effect of cultivar and development stage on infection frequency. Euphytica 26:249-255.

Parlevliet JE, Van Der Beek JG, Pieters R (1981) Presence in Morocco of brown rust, *Puccinia hordei*, with a wide range of virulence to barley. Cereal Rusts Bull 9:3–8

Peterson RF, Campbell AB,Hannah AE (1948) A diagrammatic scale for estimating rust intensity of leaves and stem of cereals. Can J Res 26: 496–500

Qualset C (1975) Sampling germplasm in a center of diversity: an example of disease resistance in Ethiopian barley. Crop genetic resources for today and tomorrow 81-96

R Core Team (2018) R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria, https://www.R-project.org/

Rehman S, Amouzoune M, Hiddar H, Aberkane H, Benkirane R, Filali-Maltouf A, Al-Jaboobi M, Acqbouch L, Tsivelikas A, Verma R, Kehel Z, & Birouk A, & Amri A (2020) Traits discovery in *Hordeum vulgare sbsp. spontaneum* accessions and in lines derived from interspecific crosses with wild *Hordeum* species for enhancing barley breeding efforts. Crop Sci 1-15. DOI: 10.1002/csc2.20360

Russell JR, Baum M, Ceccarelli S, Close TM, Grando S, Hayes PM, Matus I, Marshall DF, Del Poza A, von Korff Schmising M, Waugh R (2008) Genomic dissection of tolerance to drought stress in wild barley. 10th International Barley Genetics Symposium, Alexandria, Egypt, 5–10 April 2008

Sandhu KS, Singh D, Park RF (2014) Characterising seedling and adult plant resistance to *Puccinia hordei* in *Hordeum vulgare.* Ann Appl Biol 165:117-129

Singh R (1992) Genetic association of leaf rust resistance gene Lr34 with adult plant resistance to stripe rust in bread wheat. Phytopathology 82:835–838

Singh D, Dracatos P, Derevnina L, Zhou MX, Park RF (2015) *Rph23*: A new designated additive adult plant resistance gene to leaf rust in barley on chromosome 7H. Plant Breed. 134:62-69

Stakman EC, Stewart DM, Loegering WQ (1962) Identification of physiologic races of *Puccinia graminis* var. *tritici*. USDA Agric Res Serv E617:5–53

Steffenson BJ, Jin Y, Griffey CA (1993) Pathotypes of *Puccinia hordei* with virulence for the barley leaf rust resistance gene *Rph7* in the United States. Plant Dis. 77:867–69

Street K, Mackay M, Zuev E, Kaur N, El Bouhssini M, Konopka J, Mitrofanova O (2008) Diving into the genepool: a rational system to access specific traits from large germplasm collections. In: (Appels R, Eastwood R,

713 Lagudah E, Langridge P, Mackay M, Mcintyre L, Sharp P(eds) Proceedings of the 11th International Wheat
714 GeneticsSymposium , pp 28–31. Brisbane, Australia: Sydney University Press
715 Stubbs RW, Prescot JM, Saari EE, Dubi HJ (1986) Cereal Disease Methodology Manual. Centro Internacional de
716 Mejoramiento de Maizey Trigo (CIMMYT), Mexico, Pages: 46
717 Stuthman DD, KJ, Leonard J, Miller-Garvin (2007) Breeding Crops for Durable Resistance to Disease. Advances
718 in Agronomy, Academic Press 95: 319-367
719 Sun Y (2007) Study of *Puccinia hordei* and its host resistances in *Hordeum vulgare*. PhD Thesis, North Dakota
720 State Univ, Fargo, ND
721 Upadhyaya HD, Oritz R. (2001) A mini core subset for capturing diversity and promoting utilization of
722 chickpea genetic resources in crop improvement. Theor Appl Genet 102:1292–1298
723 USDA United States Department of Agriculture Foreign Agricultural Service (FAS), Ministry of Agriculture,
724 Rabat. (2018). (https://gain.fas.usda.gov/RecentGAIN/ Publications/ Grainand/Feed/AnnualRabatMorocco_3-
725 31-2018.pdf)
726 Venables WN, Ripley BD (2002) Modern applied statistics with S, 4th edn. Springer, New York
727 Woldeab G, Fininsa C, Singh H, Yuen J (2006) Virulence spectrum of *Puccinia hordei* in barley production
728 systems in Ethiopia. Plant pathology 55(3):351-357
729 Xu Y (2010) Plant genetic resources: Management, evaluation and enhancement. In: Molecular plant breeding.
730 Wallingford, UK, CAB International, 15–194
731 Yu X, Kong HY, Meiyalaghan V, Casonato S, Chng S, Jones EE, Butler RC, Pickering R, Johnston PA (2018)
732 Genetic mapping of a barley leaf rust resistance gene *Rph26* introgressed from *Hordeum bulbosum*. Theor Appl
733 Genet 131:2567–2580. https ://doi.org/10.1007/s0012 2-018-3173-8
734 Zadoks JC, Chang TT, Konzak CF (1974) A decimal code for the growth stages of cereals. Weed Res 14: 415-421

Figures

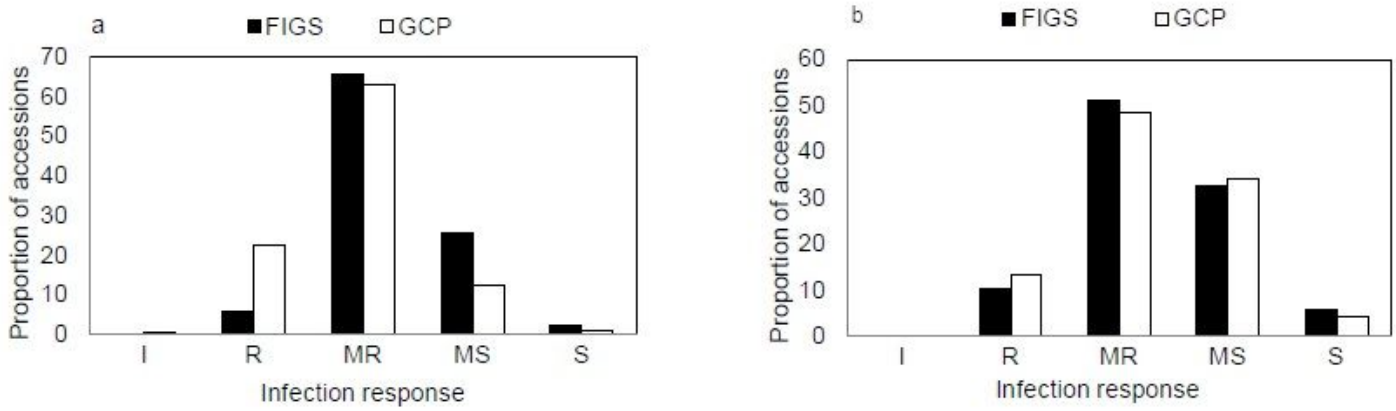

Figure 1

Frequency distributions of LR disease reactions of barley FIGS_LR and GCP subsets evaluated at the seedling stage against ISO-MRC (a) and ISO-SAT isolates (b). Here I = immune, R = resistant, MR = moderately resistant, MS = moderately susceptible, and S = susceptible.

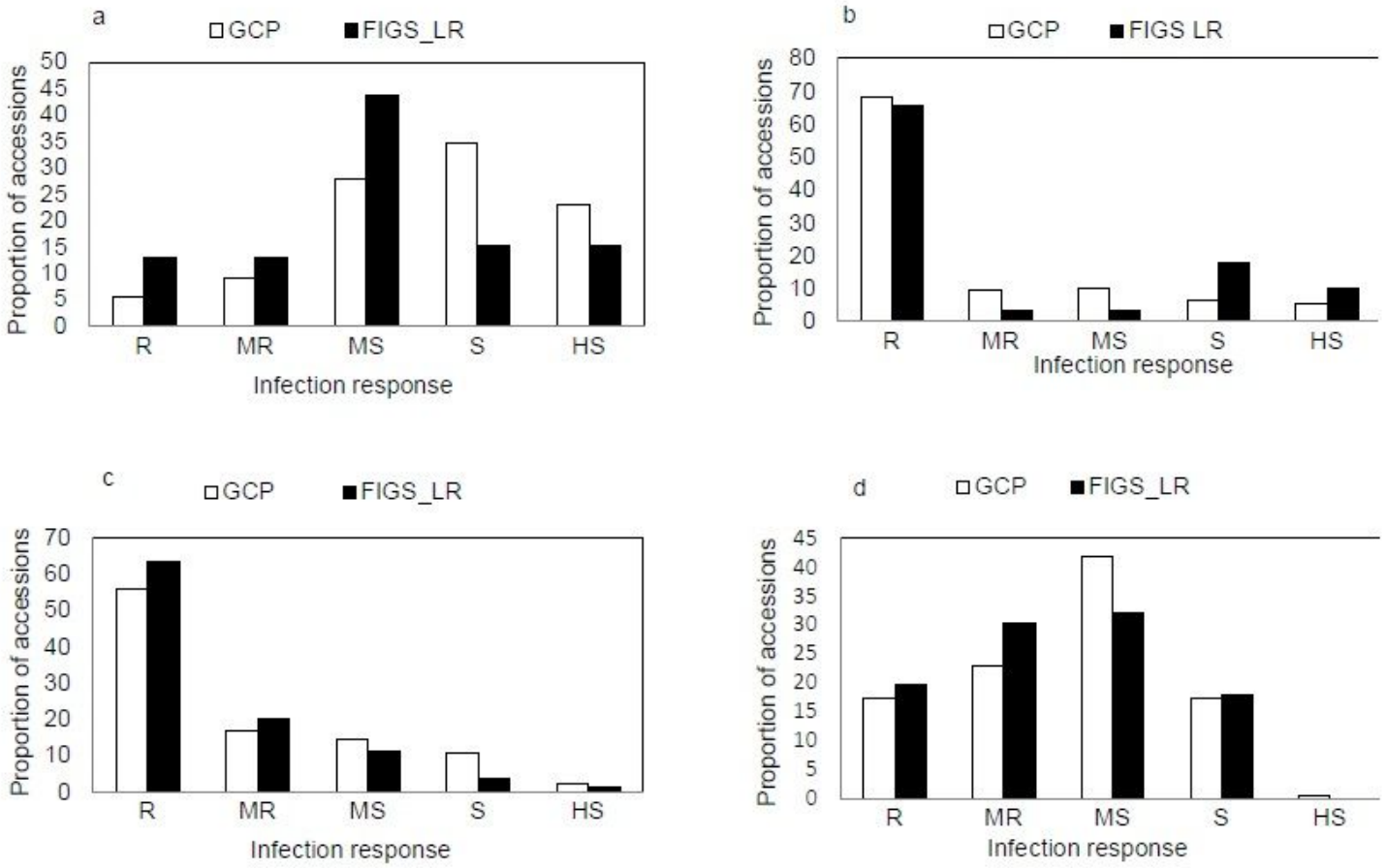

Figure 2

Frequency distributions of the adult plant resistance (APR) of FIGS_LR and GCP subsets to leaf rust under field conditions at Sidi Allal Tazi during (a) 2017 (SAT2017), (b) 2018 (SAT2018), (d) 2019 (SAT2019) cropping seasons, and at Guich during 2018 (GCH2018 (c)). Here R = resistant, MR= moderately resistant, MS = moderately susceptible, and S = susceptible.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTableS1.xlsx
- SupplementaryTableS2.xlsx